



面向下一代数据中心的软件定义存储技术研究

孙振正, 龚 靖, 段 勇, 雷俊智

(中国电信股份有限公司上海研究院 上海 200122)

摘 要:随着互联网技术的快速发展,数据的爆炸式增长,存储系统的软硬件紧耦合设计严重地限制了存储技术的发展,也越来越无法满足移动互联网和大数据时代下,对存储系统快速、多变的需求。软件定义存储(SDS)作为一种新的存储系统架构,更适合下一代数据中心的发展。从 SDS 的定义出发,对 SDS 的研究现状进行阐述,结合应用对 SDS 的需求,提出了 SDS 的参考架构,并给出了 SDS 的演进思路。

关键词:软件定义存储;软件定义数据中心;存储虚拟化;架构设计

doi: 10.3969/j.issn.1000-0801.2014.01.006

Research on Software Defined Storage Technology for the Next-Generation Data Center

Sun Zhenzheng, Gong Jing, Duan Yong, Lei Junzhi

(Shanghai Research Institute of China Telecom Co., Ltd., Shanghai 200122, China)

Abstract: With the rapid development of internet technology and the explosive growth of data, the storage system tightly coupled hardware and software design is limiting the development of storage technology severely, and unable to meet the rapidly changing needs in the mobile internet and big data era increasingly. Software defined storage as a new storage system architecture is more suitable for the development of the next-generation data centers. The SDS research status was described, started from the definition of SDS, and then SDS reference architecture combined with the SDS system requirements was proposed, finally the evolution of ideas were put forward.

Key words: software defined storage, software defined data center, storage virtualization, architecture design

1 引言

在存储业界,由于存储系统的软硬件紧耦合、管理接口不统一等限制因素,不同存储厂商间的存储设备和解决方案长期存在技术壁垒,多数现有的存储系统是单一、集成的系统,只支持特定的硬件和软件组合,存储系统缺乏灵活性,无法充分利用不断涌现的新硬件产品的能力和新平台,并且只能进行有限的扩展,无法满足用户快速增长的数据存储的需要;同时,不同存储系统之间只能实现有限程度的统一管理,由于底层不同厂商的存储设备无法形成统一、可调度的存储资源池,存储系统的管理成本开

销会日益庞大。

随着互联网技术的快速发展、云计算应用模式的快速普及以及大数据时代下数据的爆炸式增长,传统的存储系统越来越难以满足企业 IT 系统在灵活性、扩展性以及统一管理等方面的需求。为了打破传统存储系统软硬件紧耦合所造成的割裂状况,软件定义存储 (software defined storage, SDS) 应运而生。

SDS 的概念最早由 VMware 在 2012 年的 VMworld 上提出,作为 VMware 软件定义数据中心五大组成部分 (计算、存储、网络、管理和安全)中的一项,逐渐进入人们的视野。随后,EMC 公司在当年的 EMCworld 发布大会上发布



了 SDS 战略,引发了业界对 EMC 这一策略以及 SDS 的大讨论,SDS 迅速成为存储业界的研究热点。

2 软件定义存储的定义

SDS 与软件定义网络 (software defined network, SDN) 以及软件定义数据中心 (software defined data center, SDDC) 的概念颇为相似,但由于 SDS 的发展时间较短,目前业界对 SDS 并没有形成统一的定义。

IDC 通过对 SDS 市场的深入了解和研究,对 SDS 定义如下: SDS 将数据中心或者跨数据中心的各种存储资源抽象化、池化,以服务的形式提供给应用,满足应用按需(如容量、性能、QoS、SLA 等)自动化使用存储的需求。

DMTF (desktop management task force) 则认为 SDS 是由智能软件,而不是由存储硬件直接提供存储管理和自动化的基础设施。存储池的基础设施资源,在一个软件定义的存储环境中可以自动有效地分配,以符合企业的应用需求。

虽然两者的定义有所不同,但从本质上看,SDS 应如同一个生态系统,它将存储的功能从传统存储系统中抽象出来,通过软件实现,从存储系统中抽离出来单独部署,不再是硬件设备上的固件,通过软件定义,可管理来自不同厂商的所有的物理和虚拟存储资源,并按需进行自动配置。

软件定义存储应具有以下特征。

(1) 自服务

通过定义标准的应用编程接口 (API),进行存储的配置,可以满足应用程序和客户所需要的存储资源,无需人工干预。虽然这种形式的配置在公共云存储环境中(如 Amazon S3 和新兴的云计算平台 OpenStack)已经十分常见,但在企业的 IT 环境中还未实现。一些存储厂商,包括 EMC 公司、惠普和 NetApp,正逐渐提供基于 API 的配置,这种自服务特征将极大地方便应用系统管理员的配置和应用的集成,并有效地降低管理成本。

(2) 存储虚拟化

存储虚拟化可以聚合异构存储资源到一个共享的存储池,打破传统存储系统烟囱式的现状,使所有存储设备中的存储容量得以充分利用,存储虚拟化可以实现数据跨异构孤岛,在因硬件故障进行数据迁移时,实现数据的无缝迁移,并方便管理。

(3) 丰富的数据接口

SDS 可以对外提供丰富的数据接口,如文件系统接口

(NFS、CIFS 等)、块接口 (iSCSI、FC 等)、对象接口 (S3、SWIFT 等)以及可以供大数据分析的 HDFS 接口。丰富的数据接口可以为应用在存储的接口和协议方面提供更多的选择。

(4) 异构存储设备的快速接入和统一管理

SDS 的一个优势在于对异构存储设备的整合,不同类型的存储系统(如 NAS、SAN、对象存储等)以及不同厂商的存储设备可以实现快速的接入和统一管理,这为存储虚拟化提供了丰富的存储资源,也为建立全局的管理视图创造了条件。

3 软件定义存储的研究现状

“基于软件的平台会继续比基于文件或者基于对象的其他存储市场部分发展得更快,这种成长主要被跨行业 and 地域、丰富多样、密集的数据所驱动。”IDC 研究主任 Nadkarni A 在一篇声明中提到。而事实上,目前仅是软件定义存储的开端,由于传统异构存储设备之间的实现机制迥异,大存储厂商并未真正实现异构兼容,软件定义存储的研究正处于起步阶段。

3.1 研究机构的标准化研究

作为对传统存储行业的颠覆性技术,SDS 在定义、体系架构和管理接口等方面亟需一系列的标准进行规范,标准化工作对推动 SDS 的发展起着至关重要的作用,但目前,国内研究机构对 SDS 的标准化工作还没有提上日程,国际研究机构对 SDS 的标准化工作还在研究过程中,相关研究状况如下。

- IDC 在 2013 年 4 月对外正式发布文档“IDC’s Worldwide Software-Based (Software-Defined) Storage Taxonomy”,其中对 SDS 的定义进行阐述,并提出了 SDS 的评估方法以及在通用硬件上利用软件构建一个高可扩展性存储平台的解决方案。
- DMTF 在 2013 年 5 月成立了一个孵化项目 SDDC Incubator, SDS 作为其中一部分,目前正在研究过程中。
- OpenStorage 开源社区在 2013 年 10 月举办了“OpenStorage Summit 2013”大会上,组织各企业一起讨论 SDS 的定义,但还没有实质性东西输出。

3.2 厂商研究

除学术研究团体对 SDS 技术展开广泛研究外,商业公司对 SDS 技术跟进和产品研发均产生了浓厚兴趣,并进行了积极实践,但由于各家对 SDS 的理解不同,其 SDS 的系

统架构设计也有所不同。目前,对 SDS 进行深入研究的公司可以分为以下 3 类。

(1)传统的存储设备厂商(如 EMC 公司、华为技术和 NetApp)

- EMC 公司的 SDS 通过存储虚拟化平台将物理阵列中的存储(无论是基于文件、块还是对象)抽象为虚拟共享存储资源池,从而跨物理阵列实现灵活的存储使用模型,并提供应用程序和创新型数据服务。存储虚拟化平台从底层硬件阵列抽象出存储控制路径,以便可以在软件中集中执行多供应商存储基础设施的访问和管理。EMC 公司的软件定义存储系统具有控制路径和数据路径、控制路径设置和管理存储设备的策略,而数据路径执行实际读取、写入和数据服务。
- 华为技术从业务驱动存储(service-driven storage)角度理解和研究 SDS,即系统根据业务的定制化要求(主要指性能、可靠性及数据的一致性方面),将相关的控制策略下发给控制平面和数据平面,联动实现用户存储的定制化。与 EMC 公司不同的是,其将软件定义存储系统分为 3 个平面:管理平面、控制平面和数据平面。异构阵列则通过接入层统一接入数据平面,业务侧的请求统一由控制平面根据存储策略进行自动分配。
- NetApp 的 SDS 主要通过其集群操作系统 Data ONTAP 对底层的异构存储进行统一管理,并实现存储资源的策略配置以及提供可编程 API,实现应用程序自定义和工作流自动化,所有的异构存储设备通过 NetApp 设备接入。目前,NetApp 的 ONTAP 产品主要支持块和文件存储接口,对对象存储支持较弱;兼容硬件支持列表中,主要是 NetApp 认证的存储设备。

(2)虚拟化解决方案厂商(如 VMware)

作为虚拟化操作系统的龙头老大,VMware 在 SDS 相关产品中,更多地关注异构存储如何为虚拟机使用,以应用为中心的策略,可实现存储使用自动化,对异构存储池中的所有资源制定一致的策略,使存储的使用像为每个应用或虚拟机指定容量、性能和可用性要求那样简单。这种基于策略的自动化最大限度地利用了底层存储资源,同时将管理开销降至最低。数据服务(如快照、克隆和复制)作为虚拟数据服务在软件中交付,并按虚拟机进行调配和管

理,独立于底层存储硬件使得这些服务的分配极其敏捷和灵活。

(3)初创的存储软件公司(如先智)

存储功能软件化在打破传统存储设备厂商既有利益格局的过程中将发挥十分重要的作用,这也为许多初创的存储软件公司提供了参与存储产业链的机会。先智的 SDS 主要基于 OpenStack 的块存储组件 Cinder,将 SDS 架构设计为 3 个层次:接口层、管理层和存储联邦层。其中,接口层对应用提供存储管理接口及块存储接口,管理层提供资源管理、认证、统计及监控等管理功能,而存储联邦层提供异构存储设备的接入和管理。

4 软件定义存储的架构设计

4.1 系统需求

(1)实现面向应用的专业化管理策略

传统存储系统存储资源与应用独立,存储资源利用效率低,在新的存储系统设计中要做到资源整合并且针对应用进行专业化的策略管理,根据应用的变化进行弹性配额;同时,存储系统必须提供基于容器的多层次租户/应用隔离技术,通过物理隔离与权限控制相结合,实现以用户为单位、以业务为单位以及以存储容器为单位的数据隔离。

(2)系统全局自动负载均衡和智能调度

大数据时代下,新的存储系统中的物理存储主机节点规模从几万个到几十万个,多为数据密集型应用,面对超大规模的数据请求和节点数量,应高效进行节点负载均衡,充分发挥空闲节点的作用,保障高水平服务质量,提高系统运行效能。

存储系统应按照存储设备的相关特性(如存储容量、设备类型、性能指标等)为业务选择最合适的存储资源,并按照存储设备的 I/O 空闲程度等信息对应用进行基于服务质量的多层次自动负载均衡与调度功能。

(3)存储设备在线扩展和收缩

在存储设备的使用过程中,会遇到调整存储资源池的需求,要求存储资源池根据业务的需求增加或者减少存储设备。在调整的过程中,尽量减少对业务的影响,即业务不能被中断。

(4)跨数据中心的存储资源统一管理

大规模级别存储系统的构建随着数据的爆炸性增长,存储的规模越来越大。传统存储通常是在一个设备、一个



机架或一个数据中心内完成资源组织管理,而当存储容量上升到 EB 级或 ZB 级后存储很难在一个数据中心内完成。大规模的存储需要跨数据中心,跨城市、省甚至国家进行存储设备、存储数据、存储服务的组织和管理,并支持跨域的访问、备份、容灾等功能。

4.2 参考架构设计

基于以上需求分析,结合软件定义存储的特征,笔者认为在设计软件定义存储的系统时,需要着重考虑软硬件的解耦、灵活的存储配置策略、多样化的存储能力供给以及存储资源的弹性扩展等因素,软件定义存储的参考架构如图 1 所示。

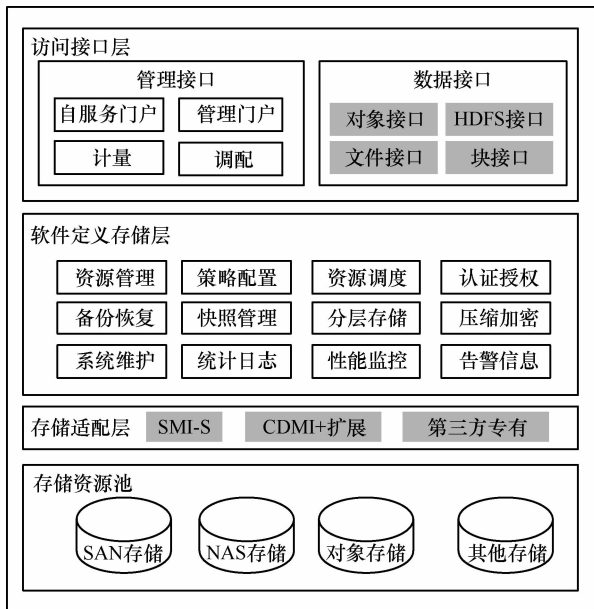


图 1 软件定义存储参考架构

软件定义存储系统主要由以下 4 层组成:访问接口层、软件定义存储层、存储适配层以及存储资源池,每一层扮演着不同的角色,功能也有所不同。

(1) 访问接口层

此层包括管理接口和数据接口,管理接口主要提供与自服务门户、管理门户以及外部计费系统的管理接口;而数据接口则通过协议转换的方式,为外部应用提供丰富的数据接口,如对象接口、文件接口、HDFS 接口以及块接口等。

(2) 软件定义存储层

此层是软件定义存储系统中十分重要的一层,通过软件对底层存储资源进行池化和统一管理,并基于存储设备的性能、可靠性等指标进行相关策略配置,为不同的应用

选择合适的存储资源;同时,此层通过软件定义可以实现或者提升传统存储设备的相关功能,如去重、快照、数据复制以及压缩加密等。

(3) 存储适配层

存储适配层主要通过多种访问接口实现对底层异构存储设备之间的访问。目前,可以通过全球网络存储工业协会(Storage Networking Industry Association, SNIA)定义的 SMI-S 协议和 CDMI 实现对存储设备的访问,但异构存储设备之间的互通性需要标准化组织的进一步推进;同时,存储适配层也支持第三方专有的访问接口,通过专有的访问接口,存储设备的相关高级功能特性能够被充分地发挥。

(4) 存储资源池

存储资源池中涵盖了业界主流的存储类型:高性能的 SAN 存储、高扩展性的 NAS 存储以及兼具二者特性的对象存储,这些存储设备都作为软件定义存储资源池中的组成部分,由上层的软件定义存储层进行统一管理和统一调度,通过访问接口层中的数据接口,为外部的应用提供存储服务。

5 软件定义存储的演进思路

SDS 的实现并非一步到位,需要分以下 3 个阶段。

阶段一:存储资源虚拟化,简化存储配置,降低管理成本。

此阶段需要结合应用系统对存储的需求,分析目前存储资源的使用状况,调整未来的购置计划,进行存储资源的合理配置,对企业内部的存储资源按照性能、可靠性等进行分级,通过存储虚拟化手段,将企业内部已有的跨越多个厂商的异构存储、DAS 资源、SSD 资源整合为统一的存储资源池,以便于应用调用;同时,通过定义相关 API,简化存储配置,实现应用系统直接访问存储资源。

存储资源全局管理视图的建立将存储管理员从繁琐的存储配置琐事中解脱出来,可以集中管理后端基础设施,通过在容量、性能、可靠性等方面对存储资源的组织,可以更加有效地满足应用系统对存储资源的需求。

阶段二:存储功能软件化,实现软硬件解耦,降低设备采购成本。

此阶段主要在统一的存储资源池上,构建完善的存储软件层,将传统存储的功能特性(如快照、精简配置以及去重等)提升到软件层次,存储功能的实现将不依赖于厂商自身的存储设备,以实现存储特性与存储硬件的解耦;同

时,存储功能软件化将支持第三方软件的无缝接入,支持用户的自定义开发,更多的存储功能创新通过软件定义的方式出现。

阶段三:基于通用硬件,提升软件定义存储的普适性。

在通用硬件的基础上,存储功能完全软件化,文件、块和对象存储设备将在通用的CPU、内存、硬盘驱动器上按需创建、灵活扩展。软件定义存储功能的不断完善和成熟,使通用硬件构建底层存储资源池成为可能,基于通用的硬件,搭载纯软件的产品将提供更大的灵活性。

6 结束语

2011年8月网景浏览器的创始人Andreessen M在华尔街日报上发表了题为“为什么软件正吞噬整个世界”的文章,拉开了“软件主世”的序幕,SDN等新概念横空出世和日渐成熟,让大家看到了下一代数据中心的雏形,也让业界对SDS未来的发展充满了信心。然而,SDS的起步稍晚,标准化的研究以及产品的研发尚处在起步阶段,再加上传统存储厂商由于各自利益,对SDS的发展战略尚不明晰,SDS的发展仍然需要一定的时间。但随着移动互联网和大数据时代的来临,作为下一代数据中心的关键能力——软件定义存储会更加灵活、高效,也将会不断向成熟、标准化和可实现的方向发展。

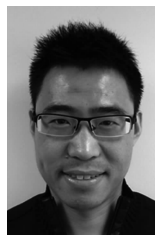
参考文献

- 1 Nadkarni A, DuBois L, Sheppard E. IDC's worldwide software-based (software-defined) storage taxonomy. <http://www.idc.com/getdoc.jsp?containerId=240500>, 2013
- 2 Baltazar H. Software-defined storage will sound the death knell for traditional storage provisioning. <http://www.netapp.com/us/system/pdf-reader.aspx?pdfuri=/tcn:10-115115-16&mc=wp-software-defined-storage-forrester.pdf>, 2013
- 3 Fichera R, Washburn D. The software-defined data center is the future of infrastructure architecture. <http://www.forrester.com/The+SoftwareDefined+Data+Center+Is+The+Future+Of+Infrastructure+Architecture/fulltext/-/E-RES81941>, 2012
- 4 IDC brings clarity to software-based/software-defined storage markets. <http://www.idc.com/getdoc.jsp?containerId=prUS24068713>, 2012
- 5 Lecat J. Is "software-defined" just a new way to whitewash old products <http://www.scality.com/is-software-defined-just-a-new-way-to-whitewash-old-products/>, 2012
- 6 Forrsights hardware survey Q3 2012. <http://www.forrester.com/>

Forrsights+Hardware+Survey+Q3+2012/-/E-SUS1631, 2012

- 7 Fichera R, Kindness A, Bartoletti D, *et al.* Optimize IT infrastructure around key workloads. <http://www.forrester.com/Optimize+IT+Infrastructure+Around+Key+Workloads/fulltext/-/E-RES61565>, 2012
- 8 Schreck G, Cullen A, Cahill M. Planning for Cloud 2.0. <http://www.forrester.com/Planning+For+Cloud+20/fulltext/-/E-RES59855>, 2011
- 9 Andrews C, Ross C F, Grannan M. Self-provisioning changes traditional sourcing practices. http://www.forrester.com/SelfProvisioning+Changes+Traditional+Sourcing+Practices/fulltext/-/E-RES58084?docid=58084&src=RSS_2&cm_mmc=Forrester_-_RSS_-_Document_-_12, 2011
- 10 软件定义新一代数据中心——VMworld 2012 大会侧记. <http://www.zcom.com/article/83373/>, 2012
- 11 VMware 推新 VMware VSAN 对软件定义存储进行了扩展. <http://www.d1net.com/storage/solution/233736.html>, 2013
- 12 EMC 发布全新软件定义存储平台 ViPR. <http://storage.chinabyte.com/146/12608146.shtml>, 2013
- 13 陈杰. 大数据场景下的云存储技术与应用. 中兴通讯技术, 2012(6)
- 14 Andreessen M. Why software is eating the world. <http://online.wsj.com/news/articles/SB10001424053111903480904576512250915629460>, 2011

[作者简介]



孙振正,男,中国电信股份有限公司上海研究院系统工程师,主要从事技术研发、云存储以及云安全的技术研究工作。

龚靖,女,中国电信股份有限公司上海研究院云存储产品经理,主要从事云存储产品研发工作。

段勇,男,中国电信股份有限公司上海研究院高级工程师、云存储专业负责人,主要研究方向为云存储、云安全、云计算的基础平台及应用等。

雷俊智,男,中国电信股份有限公司上海研究院高级工程师、长期从事固定及移动通信技术研究、移动业务网络研发、移动产品开发、云计算技术及业务研发工作。

(收稿日期:2013-12-26)